

Constraints-Based Optimization Methods for Modeling Large-Scale Biochemical Networks

Will Heuett
University of Colorado
Department of Physics

30 January 2006

Constraints-Based Optimization Methods for Modeling Large-Scale Biochemical Networks

Will Heuett
University of Colorado
Department of Physics

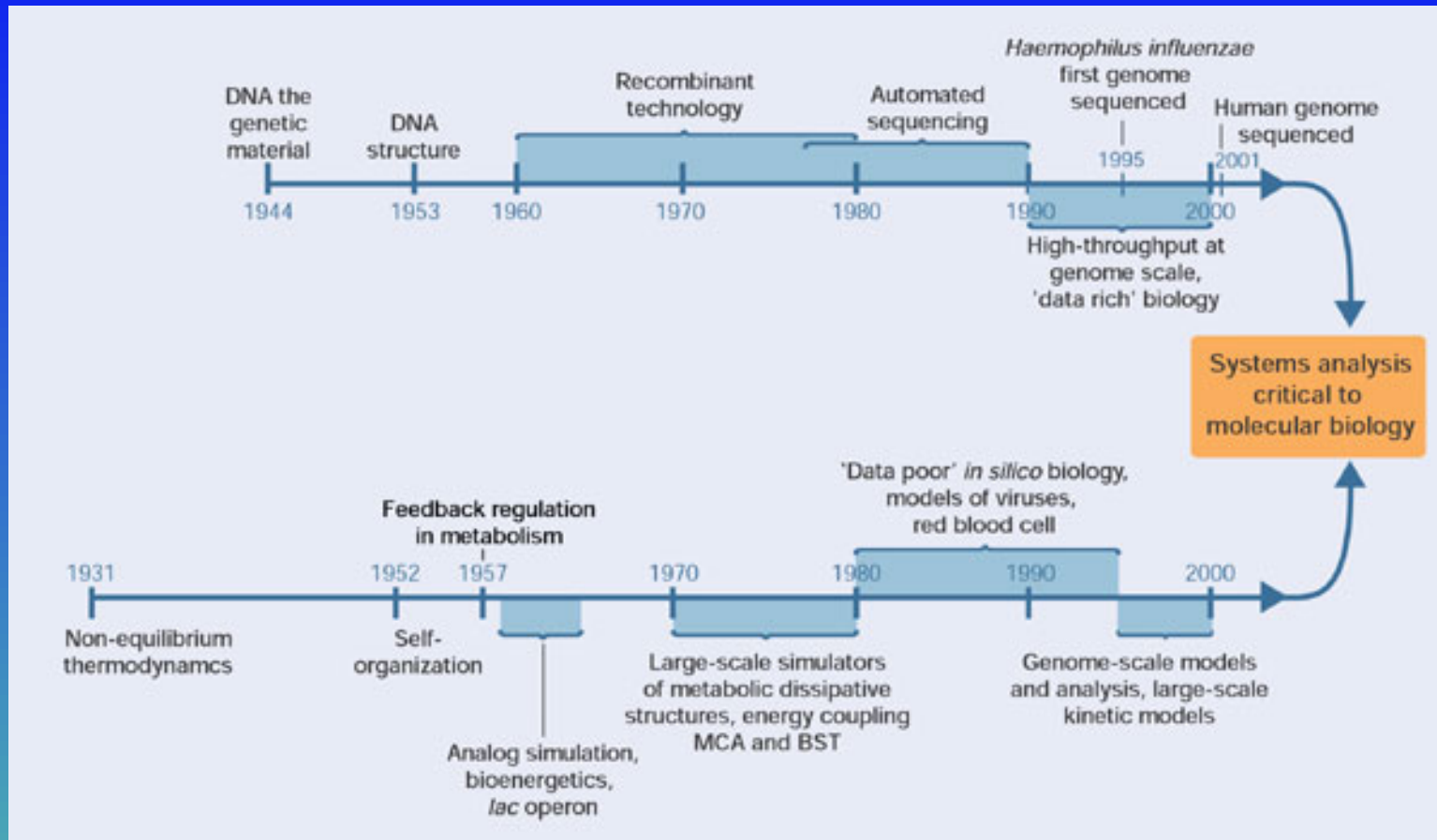
30 January 2006

I will discuss one of the leading methods used to model large metabolic reaction networks in living cells. This approach is known as the stoichiometric constraints-based approach because it uses constraints based on the stoichiometry of the system in an optimization setting.

Outline

- ★ Background and motivation for new modeling methods.
- ★ Modeling reaction kinetics using the law of mass action.
- ★ Stoichiometric constraints-based optimization approaches.
- ★ Two examples, one hypothetical and one real physical example.
- ★ Discussion.

Scaling-Up to Systems Biology



(Westerhoff and Palsson, 2004)

Classical Methods of Analysis

Unlike classical mechanics in physics, this field does not have the luxury of a long history of research because attempts to develop a general basis for a mathematical description of living organisms have only been made in recent decades.

Classical Methods of Analysis

Unlike classical mechanics in physics, this field does not have the luxury of a long history of research because attempts to develop a general basis for a mathematical description of living organisms have only been made in recent decades.

★ The Law of Mass Action:

- ★ Wilhelmy (1850) measured the velocity of mutarotation of simple sugars.
- ★ Waage and Guldberg (1864, 1867) assumed reversibility of each elementary reaction and identified the 'forward' and 'reverse' rates.
- ★ Harcourt and Esson (1866) discovered the law independently.

Classical Methods of Analysis

Unlike classical mechanics in physics, this field does not have the luxury of a long history of research because attempts to develop a general basis for a mathematical description of living organisms have only been made in recent decades.

★ The Law of Mass Action:

- ★ Wilhelmy (1850) measured the velocity of mutarotation of simple sugars.
- ★ Waage and Guldberg (1864, 1867) assumed reversibility of each elementary reaction and identified the 'forward' and 'reverse' rates.
- ★ Harcourt and Esson (1866) discovered the law independently.

★ Michaelis–Menten Enzyme Kinetics:

- ★ Named after Michaelis and Menten (1913).

Classical Methods of Analysis

Unlike classical mechanics in physics, this field does not have the luxury of a long history of research because attempts to develop a general basis for a mathematical description of living organisms have only been made in recent decades.

★ The Law of Mass Action:

- ★ Wilhelmy (1850) measured the velocity of mutarotation of simple sugars.
- ★ Waage and Guldberg (1864, 1867) assumed reversibility of each elementary reaction and identified the 'forward' and 'reverse' rates.
- ★ Harcourt and Esson (1866) discovered the law independently.

★ Michaelis–Menten Enzyme Kinetics:

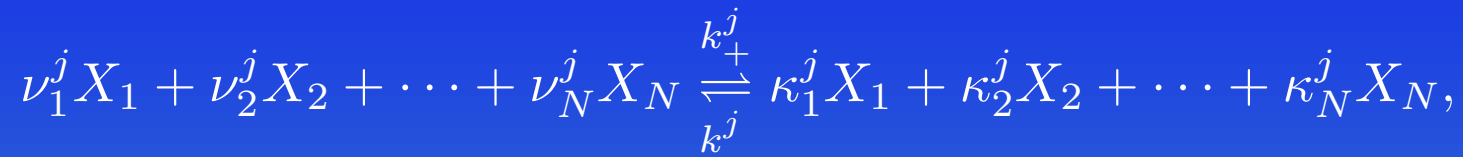
- ★ Named after Michaelis and Menten (1913).

★ Nonequilibrium Thermodynamics:

- ★ Influential work by Onsager (1931) and Hill (1989).

The Law of Mass Action

For a system involving M reactions and N chemical species with j^{th} reaction

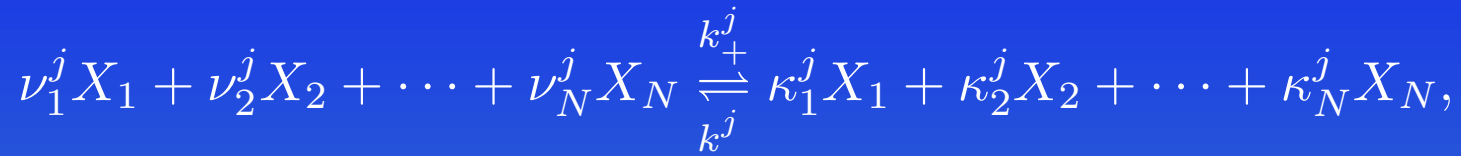


the law of mass action gives

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^M (\kappa_i^j - \nu_i^j) (k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \cdots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \cdots x_N^{\kappa_N^j}).$$

The Law of Mass Action

For a system involving M reactions and N chemical species with j^{th} reaction



the law of mass action gives

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^M (\kappa_i^j - \nu_i^j) (k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \cdots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \cdots x_N^{\kappa_N^j}).$$

A closed system will go to equilibrium, whereas an open system will go to a nonequilibrium steady state (NESS).

Detailed Balance

When in equilibrium, the forward and reverse fluxes are equal

$$k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} = k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j},$$

which yields

$$\frac{k_+^j}{k_-^j} = \frac{x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}}{x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j}} = K_{eq}^j,$$

and, for a closed loop of reactions $j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_z \rightarrow j_1$,

$$\frac{k_+^{j_1} k_+^{j_2} \dots k_+^{j_z}}{k_-^{j_1} k_-^{j_2} \dots k_-^{j_z}} = 1.$$

Open, Living Systems

Starting with the original mass-action kinetics

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^M (\kappa_i^j - \nu_i^j) (k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}),$$

the detailed balance conditions can be broken by incorporating external input and output fluxes

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^M (\kappa_i^j - \nu_i^j) (k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}) + J_i^{ext},$$

or concentration clamping

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^M (\kappa_i^j - \nu_i^j) (k_+^j c_0^{\nu_0^j} c_{N+1}^{\nu_{N+1}^j} x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j c_0^{\kappa_0^j} c_{N+1}^{\kappa_{N+1}^j} x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}).$$

Michaelis–Menten Enzyme Kinetics

For enzyme-catalyzed reactions, represented as



we can make the quasi-steady-state assumption that

$$\frac{d}{dt}(se) \approx 0 \quad \Rightarrow \quad (se) = \frac{k_{+}^1 s \cdot e}{k_{-}^1 + k_{+}^2} = \frac{s \cdot e}{K_{M,s}},$$

where $K_{M,s}$ is known as the Michaelis–Menten rate constant.

Michaelis–Menten Enzyme Kinetics

For enzyme-catalyzed reactions, represented as



we can make the quasi-steady-state assumption that

$$\frac{d}{dt}(se) \approx 0 \quad \Rightarrow \quad (se) = \frac{k_{+}^1 s \cdot e}{k_{-}^1 + k_{+}^2} = \frac{s \cdot e}{K_{M,s}},$$

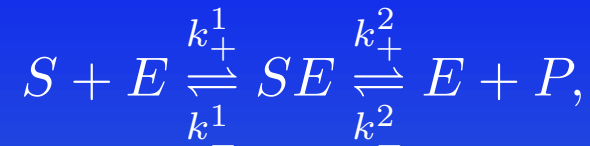
where $K_{M,s}$ is known as the Michaelis–Menten rate constant.

This yields

$$\frac{ds}{dt} = -\frac{dp}{dt} = -\frac{k_{+}^2 e_0 \cdot s}{K_{M,s} + s} = -\frac{V_{max}^+ s}{K_{M,s} + s}.$$

Reversible Michaelis–Menten Enzyme Kinetics

For



we get

$$\frac{ds}{dt} = -\frac{dp}{dt} = -\frac{V_{max}^{+} \frac{s}{K_{M,s}} - V_{max}^{-} \frac{p}{K_{M,p}}}{1 + \frac{s}{K_{M,s}} + \frac{p}{K_{M,p}}},$$

where

$$K_{M,s} = \frac{k_{-}^1 + k_{+}^2}{k_{+}^1} \quad \text{and} \quad K_{M,p} = \frac{k_{-}^1 + k_{+}^2}{k_{-}^2}.$$

Nonequilibrium Thermodynamics

The chemical potential of a species is given by

$$\mu_i = \mu_i^o + RT \ln x_i,$$

from which we get the reaction potential, given by

$$\Delta\mu^j = RT \ln \left(\frac{k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}}{k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j}} \right).$$

Nonequilibrium Thermodynamics

The chemical potential of a species is given by

$$\mu_i = \mu_i^o + RT \ln x_i,$$

from which we get the reaction potential, given by

$$\Delta\mu^j = RT \ln \left(\frac{k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}}{k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j}} \right).$$

It is easy to show that

$$- \left(k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j} \right) \Delta\mu^j \geq 0$$

and, for a closed loop of reactions $j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_z \rightarrow j_1$,

$$\Delta\mu^{j_1} + \Delta\mu^{j_2} + \dots + \Delta\mu^{j_z} = 0.$$

Limitations of Classical Methods

- ★ Typically, the biochemical networks of interest are very large and complex. As a result, it is extremely difficult to solve for analytic solutions of the models.

Limitations of Classical Methods

- ★ Typically, the biochemical networks of interest are very large and complex. As a result, it is extremely difficult to solve for analytic solutions of the models.
- ★ Experimentalists are limited in the amount of information they can gather and, in most cases, it is not possible to obtain detailed kinetic-rate information. Therefore, methods that avoid having to know this information are needed.

Limitations of Classical Methods

- ★ Typically, the biochemical networks of interest are very large and complex. As a result, it is extremely difficult to solve for analytic solutions of the models.
- ★ Experimentalists are limited in the amount of information they can gather and, in most cases, it is not possible to obtain detailed kinetic-rate information. Therefore, methods that avoid having to know this information are needed.
- ★ It is for these reasons that stoichiometric constraints-based approaches have been developed. These approaches use optimization methods and do not require any kinetic-rate information.

Stoichiometric Constraints-Based Approaches

- ★ Genome-scale constraints-based models have been developed to describe the functional states, or phenotypes, of many organisms (Westerhoff and Palsson, 2004).

Stoichiometric Constraints-Based Approaches

- ★ Genome-scale constraints-based models have been developed to describe the functional states, or phenotypes, of many organisms (Westerhoff and Palsson, 2004).
- ★ Stoichiometric Network Theory (SNT) uses the static, algebraic structure of biochemical networks, within which chemical “motion” must take place.

Stoichiometric Constraints-Based Approaches

- ★ Genome-scale constraints-based models have been developed to describe the functional states, or phenotypes, of many organisms (Westerhoff and Palsson, 2004).
- ★ Stoichiometric Network Theory (SNT) uses the static, algebraic structure of biochemical networks, within which chemical “motion” must take place.
- ★ This method of analysis has been successfully applied to systems such as *E. coli* (Edwards and Palsson, 2000), mitochondrial energy metabolism (Ramakrishna et al., 2001), and metabolism in hepatocyte cells (Beard and Qian, 2005).

Stoichiometric Network Theory

Returning to the general mass-action equation for a system of N species and M reactions

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^M (\kappa_i^j - \nu_i^j) (k_+^j x_1^{\nu_1^j} x_2^{\nu_2^j} \dots x_N^{\nu_N^j} - k_-^j x_1^{\kappa_1^j} x_2^{\kappa_2^j} \dots x_N^{\kappa_N^j}) + J_i^{ext},$$

we can rewrite the system of equations in matrix form as

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{J} + \mathbf{J}^{ext}.$$

Since this system is being driven by external fluxes, it will go to a NESS.

Flux Balance Analysis

In NESS, the concentrations of the chemical species are not changing and we have

$$\mathbf{S}\mathbf{J} = -\mathbf{J}^{ext},$$

which is known as the flux balance constraint of FBA. Note that this constraint is similar to Kirchhoff's current law of electrical circuit theory.

Additional constraints can be applied to the NESS fluxes such as

$$\begin{aligned} J_{lb}^j &\leq J^j \leq J_{ub}^j \quad \forall j \in \{1, 2, \dots, M\} \\ (J^{ext})_{lb}^i &\leq (J^{ext})^i \leq (J^{ext})_{ub}^i \quad \forall i \in \{1, 2, \dots, N\}. \end{aligned}$$

Energy Balance Analysis

Define μ as the N -dimensional vector of chemical potentials, then the M -dimensional vector of reaction potentials, $\Delta\mu$, is given by

$$S^T \mu = \Delta\mu.$$

We can define the nullspace matrix $K \in \mathbb{R}^{M \times (M-N-r)}$ with columns that form a basis for the nullspace of S , so that $SK = 0$. Then we have the constraint

$$K^T S^T \mu = K^T \Delta\mu = 0,$$

which is a constraint for the conservation of energy and is similar to Kirchhoff's loop or voltage law of electrical circuit theory.

Energy Balance Analysis

If we define the nonnegative forward and reverse reaction fluxes so that $\mathbf{J} = \mathbf{J}_+ - \mathbf{J}_-$, then the reaction potential is

$$\Delta\mu^j = RT \ln \left(\frac{J_-^j}{J_+^j} \right),$$

which leads us directly to the second law of thermodynamics, i.e.,

$$-J^j \Delta\mu^j = -RT \left(J_+^j - J_-^j \right) \ln \left(\frac{J_-^j}{J_+^j} \right) \geq 0.$$

Entropy must increase and the system must dissipate heat,

$$hdr = -\mathbf{J}^T \Delta\boldsymbol{\mu} > 0.$$

The Optimization Problem

$$\begin{aligned}
 & \min_{\mathbf{J}, \mathbf{J}_+, \mathbf{J}_-, \mathbf{J}^{ext}, \Delta\mu} && f(\mathbf{J}, \mathbf{J}_+, \mathbf{J}_-, \mathbf{J}^{ext}, \Delta\mu) \\
 & \text{s.t.} && \mathbf{S}\mathbf{J} + \mathbf{J}^{ext} = \mathbf{0} \\
 & && \mathbf{K}^T \Delta\mu = \mathbf{0} \\
 & && \text{diag}\left(e^{\Delta\mu/RT}\right) \mathbf{J}_+ - \mathbf{J}_- = \mathbf{0} \\
 & && \mathbf{J} - \mathbf{J}_+ + \mathbf{J}_- = \mathbf{0} \\
 & && \mathbf{J}_{lb} \leq \mathbf{J} \leq \mathbf{J}_{ub} \\
 & && \mathbf{0} \leq \mathbf{J}_+ < \infty \\
 & && \mathbf{0} \leq \mathbf{J}_- < \infty \\
 & && \mathbf{J}_{lb}^{ext} \leq \mathbf{J}^{ext} \leq \mathbf{J}_{ub}^{ext} \\
 & && \Delta\mu_{lb} \leq \Delta\mu \leq \Delta\mu_{ub}
 \end{aligned}$$

Sequential Quadratic Programming

- ★ We can solve the problem for any given, smooth, linear or nonlinear, objective function using a Sequential Quadratic Programming (SQP) algorithm.
- ★ The basic idea of an SQP method is to step toward an optimal solution by iteratively approximating the problem by quadratic subproblems.
- ★ A simple interpretation of an SQP algorithm is to view it as an application of Newton's method to the Karush–Kuhn–Tucker optimality conditions, i.e.,

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \nabla f(\mathbf{x}^*) - \sum_{i \in \mathcal{A}(\mathbf{x}^*)} \lambda_i^* \nabla c_i(\mathbf{x}^*) = \mathbf{0}.$$

The Quadratic Subproblem

Linearizing at the current iterate \mathbf{x}_k , we get the subproblem

$$\begin{aligned} \min_{\mathbf{p}} \quad & \frac{1}{2} \mathbf{p}^T \mathbf{H}_k \mathbf{p} + \nabla f_k^T \mathbf{p} \\ \text{s.t.} \quad & \nabla c_i(\mathbf{x}_k)^T \mathbf{p} + c_i(\mathbf{x}_k) = 0, \quad i \in \{1, 2, \dots, m\} \\ & \nabla c_i(\mathbf{x}_k)^T \mathbf{p} + c_i(\mathbf{x}_k) \geq 0, \quad i \in \{m+1, \dots, n\}, \end{aligned}$$

which gives the search direction used to update the current iterate

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$$

by doing a line search.

A Hypothetical Example

Consider:



The stoichiometric matrix is:

$$\mathbf{S} = \begin{pmatrix} -1 & 2 & -1 & -1 & 0 \\ -2 & 2 & -1 & 1 & -1 \\ 1 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & 3 & 1 \end{pmatrix}$$

The nullspace matrix is:

$$\mathbf{K} = \begin{pmatrix} -0.7163 & -0.3345 \\ -0.3205 & -0.4347 \\ 0.4710 & -0.6349 \\ -0.3958 & 0.1001 \\ -0.0752 & 0.5348 \end{pmatrix}$$

and it gives:

$$\hat{k}_1 :$$

$$0.6411A + 1.0368B + 0.7163C + 0.3205D$$

$$\hat{k}_2 :$$

$$0.8693A + 0.7692B + 0.3345C + 0.4347D$$

A Hypothetical Example

Using FBA alone to maximize D output:

Case 1			
rxn	J	species	J^e
1	0	A	1
2	0	B	0
3	1	C	0
4	0	D	-1
5	-1		

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ -\infty \end{pmatrix} \leq \mathbf{J}^e \leq \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Case 2			
rxn	J	species	J^e
1	0.05	A	10
2	10.09	B	20
3	20.17	C	20
4	9.96	D	-90
5	29.87		

$$\begin{pmatrix} 10 \\ 1 \\ 1 \\ -\infty \end{pmatrix} \leq \mathbf{J}^e \leq \begin{pmatrix} 10 \\ 20 \\ 20 \\ 0 \end{pmatrix}$$

A Hypothetical Example

Using FBA and EBA to maximize D output and minimize total energy:

rxn	J	J_+	J_-	$\Delta\mu$	J^e	species
1	-0.95	13.63	14.58	0.067	10	A
2	8.35	56.04	47.69	-0.161	20	B
3	16.94	113.85	96.91	-0.161	20	C
4	10.70	85.22	74.52	-0.134	-90	D
5	32.36	143.66	111.30	-0.255	$hdr = 13.83$	

$$f = J_D^e + \frac{\Delta\mu^T \Delta\mu}{2}$$

$$\begin{pmatrix} 10 \\ 1 \\ 1 \\ -\infty \end{pmatrix} \leq \mathbf{J}^e \leq \begin{pmatrix} 10 \\ 20 \\ 20 \\ 0 \end{pmatrix}$$

A Hypothetical Example

Using FBA, EBA, and heat constraint to maximize D and minimize total energy:

rxn	J	J_+	J_-	$\Delta\mu$	J^e	species
1	0.00	6.04	6.04	0.00	10	A
2	10.44	12.87	2.43	-1.67	20	B
3	21.33	22.12	0.79	-3.33	20	C
4	9.56	11.78	2.23	-1.67	-90	D
5	29.11	29.31	0.20	-5.00	$hdr = 250$	

With the additional constraint:

$$hdr \geq 250$$

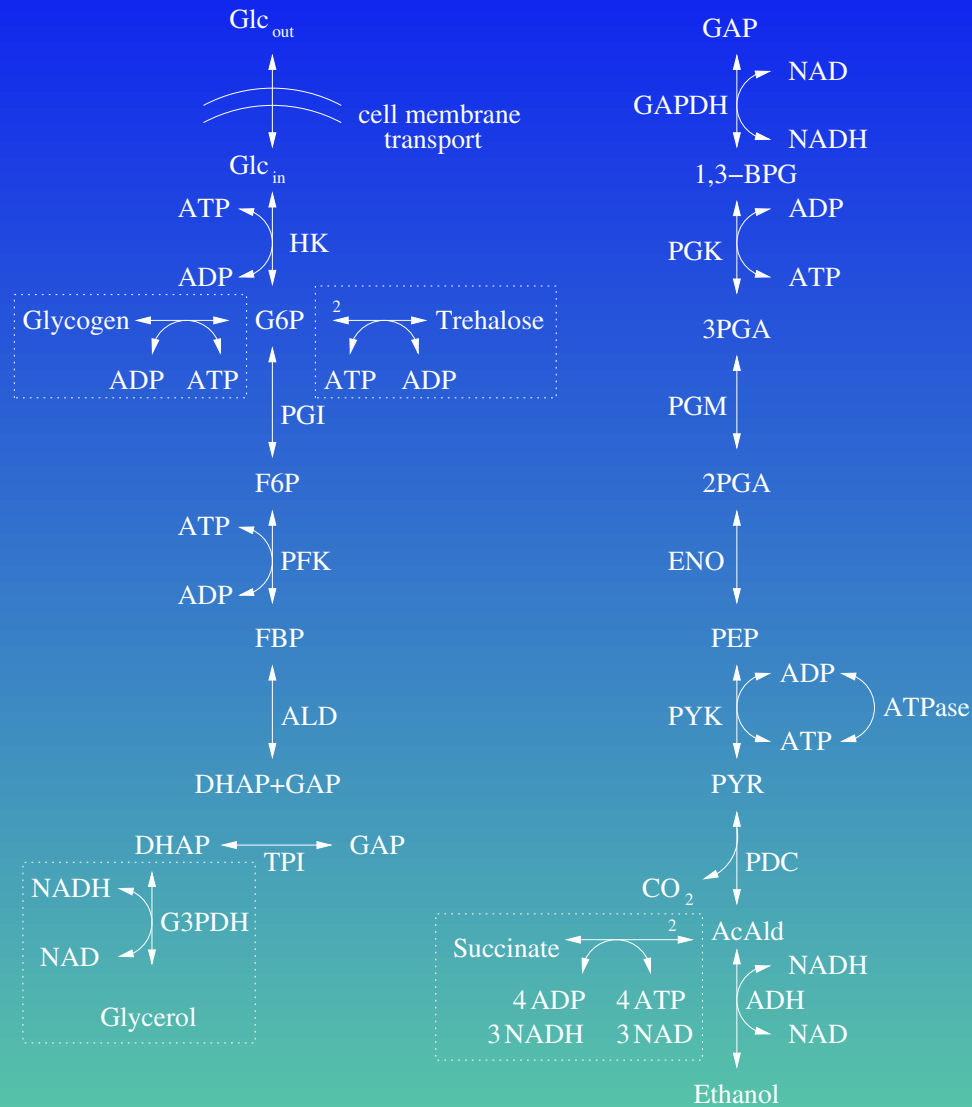
Saccharomyces cerevisiae

Overall, the net reaction of fermentation is the conversion of glucose to ethanol and carbon dioxide



- ★ Under anaerobic conditions, most of the energy from the sugar is transferred to ethanol and growth of the yeast cells is minimized.
- ★ Temperature is an important environmental factor for yeast because above the optimal temperature of 30°C, metabolism begins to slow, and when heat begins to denature the proteins in the cell, metabolism decreases rapidly.
- ★ Under anaerobic conditions with a complex medium and glucose as the substrate, a continuous culture of *S. cerevisiae* has a specific rate of heat production of $0.2 \text{ W} \cdot \text{g}^{-1}$.

Saccharomyces cerevisiae



Saccharomyces cerevisiae

Assume we measure reaction fluxes in units of $\mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg protein}^{-1}$ and reaction potentials in units of $\text{J} \cdot \text{mol}^{-1}$. Furthermore, assume that the temperature of the system is 30°C .

Experimental values are available for some reaction fluxes:

rxn	transport	HK	PGI	PFK	ALD	TPI	GAPDH
V_{min}							-24.3
V_{max}	0.36	0.84	1.26	0.68	1.19	8.4	4.4
rxn	PGK	PGM	ENO	PYK	PDC	ADH	
V_{min}	-4.8					-3.0	
V_{max}		9.4	1.35	4.05	0.65		

(Teusink et al., 2000)

$$hdr \geq 0.2W \cdot g^{-1}$$

	unbranched					branched				
	FBA	FBA, EBA, & heat				FBA	FBA, EBA, & heat			
rxn	J	J	J ₊	J ₋	$\Delta\mu$	J	J	J ₊	J ₋	$\Delta\mu$
transport	0.33	0.33	1.02	0.69	-971	0.36	0.36	0.75	0.39	-1659
HK	0.33	0.33	1.02	0.69	-971	0.36	0.36	0.75	0.39	-1658
PGI	0.33	0.33	1.02	0.69	-971	0.30	0.30	0.70	0.40	-1420
PFK	0.33	0.33	1.02	0.69	-971	0.30	0.30	0.70	0.40	-1422
ALD	0.33	0.33	1.02	0.69	-971	0.30	0.30	0.70	0.40	-1420
TPI	0.33	0.33	1.02	0.69	-971	0.23	0.23	0.64	0.41	-1126
GAPDH	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2379
PGK	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2379
PGM	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2380
ENO	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2378
PYK	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2377
PDC	0.65	0.65	1.21	0.56	-1943	0.54	0.53	0.87	0.34	-2377
ADH	0.65	0.65	1.21	0.56	-1943	0.51	0.51	0.85	0.35	-2270
ATPase	0.65	0.65	1.21	0.56	-1943	0.32	0.31	0.72	0.41	-1422
Glycogen	-	-	-	-	-	0.02	0.02	0.25	0.23	-234
Trehalose	-	-	-	-	-	0.02	0.02	0.23	0.21	-203
Glycerol	-	-	-	-	-	0.07	0.07	0.41	0.34	-476
Succinate	-	-	-	-	-	0.01	0.01	0.22	0.20	-166

Discussion

- ★ By combining FBA and EBA constraints, we are certain that the feasible solutions are mass balanced and thermodynamically realistic.
- ★ Using an SQP to solve the optimization problem allows us to combine the FBA and EBA constraints and consider many different objective functions.
- ★ This method allows us to study a system on the whole genome scale and do *in silico* experiments instead of *in vitro* or *in vivo* experiments.
- ★ The classical methods for modeling biochemical networks are limited in their power. Using stoichiometric constraints-based approaches, we are able to quantitatively study the possible phenotypes of a system.
- ★ SNT has been shown to be a very accurate and useful tool for studying mutant and disease affected organisms.

Acknowledgments

- ★ Hong Qian, my advisor.
- ★ James Burke, Mark Kot, and Dan Beard.
- ★ UW Department of Applied Math.
- ★ NSF and DoD.